

## Hasil Tagging Berdasarkan Ketiga Metode

Kata	Tag	Predicted Tag
Menteri	NNP	NNP
pertahanan	NNP	NN
AS	NNP	NNP
dijadwalkan	VB	VB
mengunjungi	VB	VB
India	NNP	NNP
.	Z	Z
Tata	NNP	NNP
Power	NNP	NNP
menyuplai	VB	VB
batu bara	NN	NN
pada	IN	IN
tahun	NN	NN
2000	CD	CD
.	Z	Z
Pemerintah	NN	NN
hati-hati	JJ	JJ
dalam	IN	IN
mengelola	VB	VB
BUMN	NN	NN
.	Z	Z
Perusahaan	NN	NN
baru	JJ	JJ
tersebut	PR	PR
mencanangkan	VB	VB
target	NN	NN
perolehan	NN	NN
laba bersih	NN	NN
.	Z	Z
Menteri	NNP	NNP
pertahanan	NNP	NN
mengunjungi	VB	VB
pangkalan	NN	NN
udara	NN	NN
.	Z	Z

Kata	Tag	Predicted Tag
Menteri	NNP	NNP
pertahanan	NNP	NNP
AS	NNP	NNP
dijadwalkan	VB	VB
mengunjungi	VB	VB
India	NNP	NNP
.	Z	Z
Tata	NNP	CD
Power	NNP	NNP
menyuplai	VB	VB
batu bara	NN	NN
pada	IN	IN
tahun	NN	NN
2000	CD	CD
.	Z	Z
Pemerintah	NN	NN
hati-hati	JJ	NNP
dalam	IN	IN
mengelola	VB	VB
BUMN	NN	NNP
.	Z	Z
Perusahaan	NN	NN
baru	JJ	CC
tersebut	PR	PR
mencanangkan	VB	VB
target	NN	NN
perolehan	NN	NN
laba bersih	NN	MD
.	Z	Z
Menteri	NNP	NNP
pertahanan	NNP	NNP
mengunjungi	VB	VB
pangkalan	NN	NN
udara	NN	NNP
.	Z	Z

Kata	Tag	Predicted Tag
Menteri	NNP	NNP
pertahanan	NNP	NNP
AS	NNP	NNP
dijadwalkan	VB	VB
mengunjungi	VB	VB
India	NNP	NNP
.	Z	Z
Tata	NNP	NNP
Power	NNP	NNP
menyuplai	VB	VB
batu bara	NN	NN
pada	IN	IN
tahun	NN	NN
2000	CD	CD
.	Z	Z
Pemerintah	NN	NN
hati-hati	JJ	JJ
dalam	IN	IN
mengelola	VB	VB
BUMN	NN	NN
.	Z	Z
Perusahaan	NN	NNP
baru	JJ	JJ
tersebut	PR	PR
mencanangkan	VB	VB
target	NN	NN
perolehan	NN	NN
laba bersih	NN	NN
.	Z	Z
Menteri	NNP	NNP
pertahanan	NNP	NNP
mengunjungi	VB	VB
pangkalan	NN	NN
udara	NN	NN
.	Z	Z

Urutan hasil metode dari kiri adalah baseline, classification dan HMM. Diatas adalah merupakan hasil dari contoh 3 kalimat pertama dari kalimat test.

## Perbandingan Akurasi Ketiga Metode

```
In [19]: 1 accuracyBaseline = 0
          2 for i in range (len(dfPredictedBaseline)):
          3     if dfPredictedBaseline.iloc[i,1] == dfPredictedBaseline.iloc[i,2]:
          4         accuracyBaseline += 1
          5 accuracyBaseline = accuracyBaseline / len(dfPredictedBaseline)
          6 print("Akurasi: ", accuracyBaseline)
```

Akurasi: 0.948051948051948

```
In [13]: 1 accuracyClassification = 0
          2 for i in range (len(dfPredictedClassification)):
          3     if dfPredictedClassification.iloc[i,1] == dfPredictedClassification.iloc[i,2]:
          4         accuracyClassification += 1
          5 accuracyClassification = accuracyClassification / len(dfPredictedClassification)
          6 print("Akurasi: ", accuracyClassification)
```

Akurasi: 0.8311688311688312

```
In [27]: 1 accuracyHMM = 0
2 for i in range (len(dfPredictedHMM)):
3     if dfPredictedHMM.iloc[i,1] == dfPredictedHMM.iloc[i,2]:
4         accuracyHMM += 1
5 accuracyHMM = accuracyHMM / len(dfPredictedHMM)
6 print("Akurasi: ", accuracyHMM)
```

Akurasi: 0.8961038961038961

Dapat dilihat dari akurasi yang telah dihasilkan model Baseline mendapatkan nilai paling besar yaitu diatas 90%. Hal ini dikarenakan metode baseline menggunakan pendekatan dimana:

- Jika kata pada data uji terdapat pada latih maka, tag yang dihasilkan adalah tag terbanyak yang dihasilkan dari kata tersebut.
- Dan jika kata tidak terdapat pada data uji maka tag yang dihasilkan adalah tag yang terbanyak keluar dari keseluruhan data dalam kasus ini tag (NN) merupakan keluaran terbanyak.

Kemudian disusul oleh model HMM, model HMM mendapatkan posisi ke-2 dengan nilai akurasi 89% dikarenakan metode tersebut akan melakukan perhitungan kedepan dan kebelakang dimana hal ini membuat jalur pada jaringan saraf tiruannya untuk membuat jalur yang paling optimal. Dikarenakan urutan dari suatu kalimat juga sangat menentukan hasil akhir dari model HMM. Hal ini juga menjadi nilai tambah dikarenakan ketika setiap kata dan tag memiliki suatu probabilitas terhadap suatu kalimat, maka tag dari suatu kata akan saling bergantung satu sama lainnya.

Diposisi paling akhir yaitu metode classification, classification yang dibangun belum menentukan urutan dari suatu kalimat sehingga akurasi yang didapatkan terlihat kecil. Dan model ini juga masih harus diberikan aturan-aturan khusus, sehingga asumsi yang dilakukan oleh model ini berdasarkan aturan-aturan yang telah dibangun. Hal ini menjelaskan kenapa classification mendapatkan nilai yang paling kecil dari ketiga model yang telah dibangun.